THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 258, No. 19, Issue of October 10, pp. 11864–11870, 1983
Printed in U.S.A.

Chicken Calmodulin Genes

A SPECIES COMPARISON OF cDNA SEQUENCES AND ISOLATION OF A GENOMIC CLONE*

(Received for publication, May 23, 1983)

John A. Putkey‡, Kuei Fang Ts'ui§, Toshio Tanaka, Lisette Lagacé, Joseph P. Stein¶, Eugene C. Lai, and Anthony R. Means

From the Department of Cell Biology, Baylor College of Medicine, Houston, Texas 77030

A cDNA library, prepared from poly(A⁺) mRNA isolated from chicken brain, was screened for calmodulin sequences using the cloned full length structural gene from *Electrophorus electricus* as probe (Lagacé, L., Chandra, T., Woo, S. L. C., and Means, A. R. (1983) *J. Biol. Chem.* 258, 1684–1688). Fifteen positive signals were detected among 4500 recombinant clones from which two overlapping clones (pCB12 and pCB15) were selected for subsequent DNA sequencing. The combined unique sequences of the two cDNA clones yielded 1395 base pairs and contained the entire coding region for calmodulin, 94 base pairs of the 5′-nontranslated region, and the entire 3′-nontranslated region of 857 base pairs. The derived amino acid sequence of chicken calmodulin is identical with that of the bovine or human protein. Compared to the eel, there is a single conservative amino acid substitution at position 74 which is occupied by Arg in the chicken and Lys in the eel. The overall nucleotide homology between the amino acid coding regions of chicken and eel calmodulin mRNA is 79%. However, the 5′- and 3′-nontranslated regions of the chicken and eel mRNA for calmodulin are highly diverged with sequence homologies of 21 and 29%, respectively. The cDNA clones were used as probes to determine the size and distribution of calmodulin mRNA in a variety of chicken tissues. In all tissues examined, two species of mRNA for calmodulin were detected at 1600 and 1900 nucleotides. Both mRNAs occurred in the cytoplasm with an abundance ratio of 4:1 for the 1600 and 1900 species, respectively. The two mRNAs appear to result from differential processing of transcripts from a single calmodulin gene. Screening of a chicken genomic phage library using pCB12 as a probe yielded a single positive designated CL-1 which contains a DNA insert of 13.5 kilobase pairs. Partial sequencing of CL-1 has confirmed the presence of sequences which code for calmodulin. A comparison of the restriction maps of CL-1 and pCB12 and pCB15 indicates that CL-1 contains at least 3 intervening sequences.

CaM[1] is a calcium-binding protein ($M_r \approx 16{,}700$) which functions in a regulatory capacity for a diverse array of enzymatic reactions and biochemical pathways. The pleiotropic nature of CaM-regulated events has sparked intensive research efforts concerning the biology, structure, and mechanism of action of CaM (for review see Refs. 1–3). This protein is constitutively expressed in target cells for both steroid and peptide hormones (3), but the intracellular concentration is elevated in transformed cells (4) as well as at the G_1/S boundary of the growth cycle of mammalian cells in culture (5). In order to evaluate the regulatory mechanisms that are operable under these circumstances, this laboratory initiated a series of studies which have employed recombinant DNA techniques to isolate the nucleic acids involved in CaM synthesis.

In an initial report, Munjaal *et al.* (6) described the isolation of a cDNA for eel CaM (pCM109) which contained CaM-specific sequences corresponding to amino acid residues 93–148. Lagacé *et al.* (7) then used pCM109 to isolate a full length cDNA for eel CaM (pCM116) in which are present multiple functional polyadenylation sites in the 3′-nontranslated region. We now report the isolation of a cDNA clone for chicken brain CaM. The chicken cDNA has been used for three purposes: (*a*) to determine, for the first time, the amino acid sequence of chicken CaM; (*b*) to compare the structure of calmodulin cDNAs from two species (chicken and eel); and (*c*) to use the cDNA as a probe to select genomic clones for CaM from a chicken gene library.

MATERIALS AND METHODS

Enzymes and Radioisotopes—All restriction endonucleases were obtained from either Bethesda Research Laboratories or New England Biolabs. S1 nuclease was obtained from Miles Laboratories, Inc. Terminal transferase and bacterial alkaline phosphatase were obtained from Bethesda Research Laboratories. T_4 polynucleotide kinase was obtained from New England Nuclear. DNase I and DNA polymerase I were obtained from Worthington and Boehringer Mannheim, respectively. [γ-^{32}P]ATP and [α-^{32}P]NTPs were obtained from Amersham Corp. Hydrazine was obtained from Pierce Chemical Co. Piperidine, from Fisher, was redistilled before use. All other chemicals were reagent grade or electrophoresis grade quality.

Isolation of mRNA and DNA—White leghorn hens were sacrificed by decapitation and their brain tissues were placed in liquid nitrogen immediately after dissection. DNA and total RNA were isolated from the frozen tissue as previously described (8). Poly(A⁺)-containing RNA was prepared from total RNA by chromatography on oligo(dT)-cellulose according to the method of Aviv and Leder (9). Nuclei-free homogenates were prepared according to the procedure of Bush (10) as modified by Tsai *et al.* (11). RNA was isolated from nuclei-free homogenates by extraction with phenol/SDS as described above (8) and chromatographed on oligo(dT)-cellulose (9) to enrich for cytoplasmic poly(A⁺)-containing RNA.

* This work was supported in part by American Cancer Society Grant ACS BC-326 and Robert A. Welch Foundation Grant Q-611. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡Supported by Cancer-directed Fellowship DRG-582 of the Damon Runyon-Walter Winchell Cancer Fund. To whom correspondence should be addressed.

§Current address, Shanghai Institute of Biochemistry, Chinese Academy of Sciences, Shanghai, Peoples Republic of China.

¶Current address, Division of Endocrinology, University of Texas, Houston, Texas 77030.

[1] The abbreviations used are: CaM, calmodulin; SDS, sodium dodecyl sulfate; dsDNA, double-stranded cDNA; kb, kilobase pairs.

Construction of a Chick Brain cDNA Library—Total poly(A$^+$) mRNA from chicken brain was used to direct the synthesis of daDNA with avian myeloblastosis virus reverse transcriptase as described previously (12). After removal of the hairpin loop by S1 nuclease digestion (13), the daDNA was sized on a Sephadex G-100 column and only the front half of the flow through peak was used for cloning. The plasmid pBR322 was cleaved with *Pst*I and terminal polynucleotide "tails" of (dC) and (dG) were added to the 3'-ends of the daDNA and cleaved plasmid, respectively (14). The tailed daDNA was then reannealed into the tailed plasmid as previously described (12).

Escherichia coli K12 strain RRI was used as host for transformation by the chimeric plasmid. The transformation was performed as previously described (15). About 6000 tetracycline-resistant (Tc^R) ampicillin-sensitive (Ap^S) colonies containing chimeric plasmid were obtained from 30 μg of mRNA.

Colony Screening by in Situ Hybridization—Tetracycline-resistant (Tc^R) and ampicillin-sensitive (Ap^S) colonies were picked onto fresh agar plates containing tetracycline and grown overnight at 37 °C overnight. Replicate filters were prepared on Whatman 541 paper and lysed *in situ* by the method of Grunstein and Hogness (16). The dried filters were prehybridized overnight at 68 °C in 6 × SSC, 1 × Denhardt's solution (17). pCM116, a full length cDNA clone for eel CaM (7), was then labeled by nick translation (18) and hybridized to the filters for 16 h at 68 °C in a solution of 6 × SSC, 1 × Denhardt's, 1 mM EDTA, 0.5% SDS. The filters were subjected to four 15-min washes at room temperature in 2 × SSC, and positive colonies were identified by autoradiography.

Individual colonies were cultured in 1.5 ml of L broth containing 10 μg/ml of tetracycline for 16 h. Bacteria were collected by centrifugation, and plasmid DNA was extracted by a rapid alkaline procedure (19) and analyzed by restriction endonuclease digestion.

Preparation of Recombinant Plasmid DNA—Bacterial colonies were grown in 1-liter cultures of M9 medium at 37 °C to an optical density of 0.6 at 600 nm. Chloramphenicol was added to 25 μg/ml, and the culture was incubated in a shaker for an additional 12 h. Clear lysates were prepared from these bacteria by the method of Katz *et al.* (20) and centrifuged in a 76% cesium chloride density gradient (21). The plasmid DNA band was collected, extracted with isoamyl alcohol, and precipitated with ethanol. The DNA was finally resuspended and stored in 10 mM Tris-HCl, pH 7.5, 1 mM EDTA (TE). Plasmid DNA to be used for DNA sequencing was passed over a Sepharose 4B column to remove contaminating RNA.

Restriction Endonuclease Digestion and Analytical Gel Electrophoresis—Restriction endonucleases were purchased from either Bethesda Research Laboratories or New England Biolabs, and digestions of DNA were carried out according to their recommended conditions. The DNA digests were analyzed by electrophoresis through either a 1% horizontal agarose gel or a 5% polyacrylamide vertical slab gel. Both gel systems utilized a buffer of 89 mM Tris-borate, pH 8.4, 2.5 mM EDTA. The DNA fragments were visualized by UV illumination after staining with ethidium bromide.

Preparation of DNA Probes—250 μg of chimeric plasmid DNA were digested with the appropriate restriction endonucleases and subjected to electrophoresis in a 1% low melting agarose (SeaPlaque, FMC Corp.) gel. The DNA fragment containing the desired insert sequences was excised from the gel and incubated at 68 °C for 10 min. The aqueous fraction was extracted with saturated phenol twice at 37 °C and twice at 4 °C. The DNA was then ethanol precipitated and resuspended for storage in TE. The DNA probes were labeled by nick translation as previously described (18).

DNA and RNA Blotting and Hybridization—DNA in agarose slab gels was transferred unidirectionally or bidirectionally onto nitrocellulose filters by the method of Southern (22). The filters were baked at 65 °C for 3 h and then prehybridized and hybridized to ^{32}P-labeled DNA probes as described above for colony screening by *in situ* hybridization. The filters were then washed several times at 68 °C in 2 × SSC, 0.5% SDS for a total of 8 h.

RNA was electrophoresed in a 1.5% agarose gel in the presence of 6% formaldehyde (23) and transferred to nitrocellulose paper (24). The immobilized RNA was pretreated and hybridized to nick-translated DNA probes. The hybridization was performed at 42 °C for 16 h in the presence of 45% formamide. The filters were washed to removed nonspecifically associated radioactivity.

DNA Sequencing—DNA fragments were treated with bacterial alkaline phosphatase and labeled at the 5' ends with T4 polynucleotide kinase in the presence of [γ-^{32}P]ATP. The labeled DNA was recut with an appropriate restriction endonuclease, and the desired fragments were isolated from a 5% polyacrylamide gel. The fragments were degraded chemically according to the method of Maxam and Gilbert (25, 26) and analyzed by electrophoresis on 0.3-mm urea-acrylamide gels according to Sanger and Coulson (27).

Autoradiography—Hybridization filters and DNA sequencing gels containing ^{32}P-labeled DNA fragments were exposed to Kodak X-OMAT AR5 or Fuji NIF-RX x-ray films in the presence or absence of Dupont Cronex Lightning Plus or Quanta III intensifying screens. The films were exposed at −20 or −70 °C.

Biosafety Precautions—All cloning experiments were carried out in compliance to the National Institutes of Health guidelines for recombinant DNA research.

RESULTS

Selection of Chicken cDNA Clones Using an Eel cDNA Probe—A cDNA library of about 6000 recombinant clones was constructed using 30 μg of chicken brain poly(A$^+$) mRNA as starting material. Screening 4500 of these clones with the ^{32}P-labeled cDNA insert from pCM116, which represents the full length cDNA for eel CaM (7), yielded 15 colonies with positive hybridization signals. These colonies were examined by the rapid restriction analysis method to determine the size and orientation of the cDNA insert. Of the 15 colonies two were chosen for subsequent sequence analysis. The first colony selected (designated pCB15) contained the largest cDNA insert of 1.3 kb. The flanking *Pst*I sites into which the cDNA was inserted were not regenerated in pCB15; however, the cDNA contained two internal *Pst*I sites. The second clone (pCB12) contained an insert of 0.7 kb. Similar to pCB15, the flanking *Pst*I insertion sites were not regenerated in pCB12. Unlike pCB15 and the other colonies which were examined by restriction analysis, pCB12 contained a single cleavage site for *Pst*I and also contained an internal *Eco*RI site which was absent in pCB15. These data suggested that these two plasmids contained overlapping cDNA inserts, an hypothesis that was subsequently confirmed by sequence analysis.

Sequence Analysis of pCB15 and pCB12—The restriction maps and sequencing strategy for pCB15 and pCB12 are shown in Fig. 1. The combined primary nucleotide sequence is shown in Fig. 2. The length of the parent mRNA is 1395 nucleotides and includes 94 nucleotides of 5'-nontranslated region and an 857-nucleotide 3'-nontranslated region. pCB12 contains the 5'-nontranslated region, all of the CaM structural sequence, and a small fraction of the 3'-nontranslated region. pCB15 contains the majority of the CaM amino acid coding region, all of the 3'-nontranslated region, including a poly(A) addition/termination signal (AATAAA). Twenty nucleotides 3' from this signal is a poly(A) tract of 28 nucleotides which presumably represents a portion of the original poly(A) tail in the mRNA. pCB12 and pCB15 overlap each other by approximately 430 nucleotides and, therefore, originated from the same mRNA species.

The nucleotide sequence which codes for the amino acids of chicken CaM extends from position 95 to 538 and represents 149 amino acids. The 5'-nontranslated region contains a termination codon (position 5 of the cDNA sequence) which is in the same reading frame as the initiation codon of CaM, thus precluding the possibility that chick CaM is translated with a signal peptide. The amino acid sequence shown in Fig. 3 is identical with the peptide sequence for bovine and human brain CaM (28, 29). The amino acid sequence for chicken CaM differs from eel CaM by a single conservative amino acid substitution of an Arg in the chicken for a Lys in the eel at amino acid position 74.

The amino acid sequence of CaM can be divided into four domains which exhibit sequence homology with one another (28). The highest degree of homology is seen between domains I/III and II/IV. Table I shows the nucleotide sequence ho-

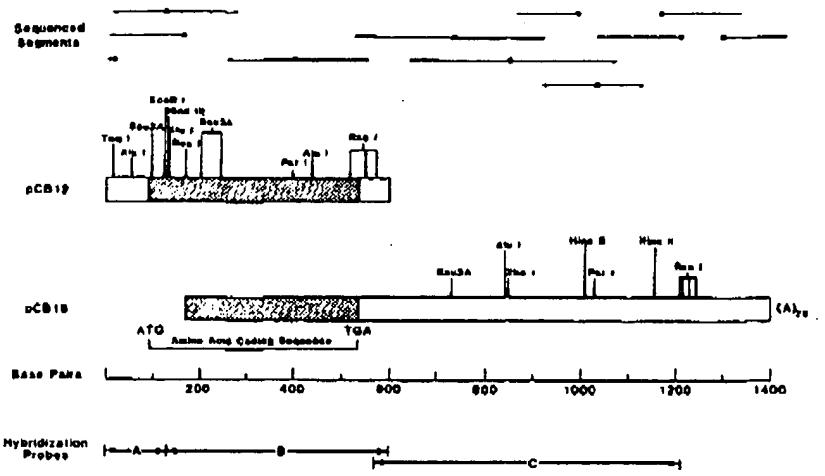

FIG. 1. **Strategy for sequencing pCB12 and pCB15.** The overlapping cDNA clones are shown together with the restriction endonuclease sites which were utilized for DNA sequencing. All restriction sites in pCB12 which are in the overlap region with pCB15 are also found in pCB15. The *open bars* represent the 3′- and 5′-nontranslated regions while the *striped bars* indicate the amino acid-coding regions. The indicated hybridization probes A, B, and C were generated by digestion of pCB12 with *Hin*dIII and *Ava*II and by digestion of pCB15 with *Rsa*I. These probes were used for detection of mRNA for CaM in chicken tissues.

mologies between those portions of the chicken cDNA clones that correspond to the amino acid domains. Also shown are the nucleotide sequence homologies between the domains of the chicken cDNA compared to those of the eel cDNA. As anticipated, the highest degree of nucleotide sequence homology parallels the interdomain amino acid sequence homologies. In addition the cross-species nucleotide sequence homology conforms to the predicted hierarchy of homology. That is, the highest degree of homology is seen between domains I/III and II/IV.

Each molecule of CaM can bind 4 molecules of calcium. The importance of calcium binding to the function of CaM implies that those portions of the molecule which participate in calcium binding will be more highly conserved than most other portions of the molecule. Based on the crystalline structure of parvalbumin (30) the putative calcium-binding sites of CaM can be assigned. One binding site exists in each of the four amino acid domains. Each site is composed of 12 amino acid residues, 6 of which serve as ligands for Ca^{2+} in the metal-protein complex. Table II lists the per cent homology, between chicken and eel, of the protein domains of CaM and their corresponding calcium-binding subdomains. Within each amino acid domain the region of putative calcium binding exhibits the highest degree of nucleotide sequence homology.

In contrast to the amino acid coding portions of the chicken and eel cDNAs, the 5′- and 3′-nontranslated regions of these cDNAs have diverged to a large extent. This is particularly true for the 5′-nontranslated region which is extremely GC-rich in the chicken.

mRNA for CaM in Chicken Tissues—Eel CaM is translated from 3 mRNAs which have distinctly different sizes (7) as a result of differential termination/polyadenylation. In light of this, it was of interest to determine the number of mRNA species which are specific for chicken CaM. For this purpose, three hybridization probes were prepared from pCB12 and pCB15 (see Fig. 1). Probe A contains the entire 5′-nontranslated region and extends to the unique *Hin*dIII site in the amino acid coding portion of the cDNA. Probe B extends from the *Hin*dIII site to the 3′-end of pCB12 and contains the majority of the amino acid coding sequences. Probe C is a 600-base pair fragment resulting from a *Rsa*I digest of pCB15 and is derived exclusively from the 3′-nontranslated region. Together these probes represent 85% of the cDNA.

The probes were radioactively labeled by nick translation and hybridized to both total and cytoplasmic chicken brain poly(A$^+$) RNA as well as total chicken muscle poly(A$^+$) RNA. As shown in Fig. 3, all three probes give identical hybridization patterns and detect two species of mRNA. The stronger hybridization signal corresponds to an mRNA of 1600 nucleotides in length, while the weaker signal corresponds to an mRNA of 1900 nucleotides in length. Both hybridization signals are observed when either total or cytoplasmic chicken brain or total muscle poly(A$^+$) RNA is used. Essentially identical results were obtained with RNA from chicken oviduct, gizzard, and liver (data not shown).

Hybridization of the three probes to brain cytoplasmic and total muscle poly(A$^+$) RNA yields several signals which are not evident in total brain poly(A$^+$) RNA. All three probes hybridize to RNA species of less than 1600 nucleotides in brain cytoplasmic poly(A$^+$) RNA. These hybridization bands are diffuse and are not evident in total brain poly(A$^+$) RNA which suggests that they result from degradation of the two larger RNA species during the additional manipulations employed to obtain cytoplasmic poly(A$^+$) RNA. The hybridization pattern seen for total muscle poly(A$^+$) RNA is intriguing in that there are two distinct hybridization signals corresponding to RNAs which are considerably larger than the primary signals of 1600 and 1900 nucleotides. These RNA species are detected with only probe A, and it is, therefore, unlikely that they are precursor forms of the 1600 and 1900 RNA species. The identity of these large poly(A$^+$) RNA species is unknown at present although the possibility exists that those nucleotide sequences in probe A, which is derived from the 5′-end of the chick CaM cDNA, recognizes RNA species which code for other proteins. Probe A contains regions which are extremely GC-rich in the 5′-nontranslated region and which could hybridize to similar GC-rich regions in other RNAs.

Analysis of Genome Complexity—Three hybridization probes were prepared for use in an analysis of genome complexity by the method of Southern (22). The hybridization probes were derived from pCB12 and pCB15 and are shown in Fig. 5 as probes A′, B′, and C′. Genomic DNA, obtained from chicken liver, was digested with three restriction endonucleases (*Bam*HI, *Eco*RI, or *Hin*dIII). The digests were resolved on a 1% agarose gel and hybridized to the three ^{32}P-labeled probes as shown in Fig. 4. Hybridization of all three probes to DNA digested with *Bam*HI (Fig. 4, *lanes B*) gave a single signal with a size of 25 kb. Similarly, hybridization of all three probes to DNA digested with *Eco*RI (Fig. 4, *lanes E*) yielded a single signal with a size of 10 kb. Hybridization to

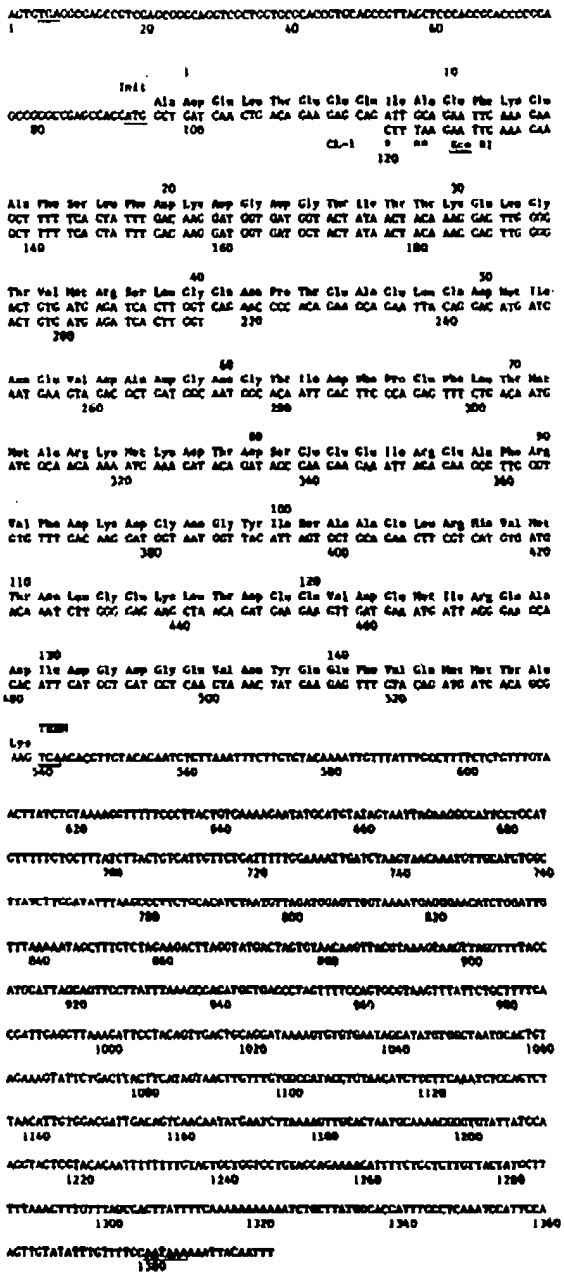

FIG. 2. **Combined nucleotide sequence of pCB12 and pCB15 and predicted amino acid sequence of chicken CaM.** The *numbers below the line* indicate nucleotide positions; the *numbers above the line* indicate amino acid position. The initiation (*Init*) and termination (*Term*) codons are *underlined* as are the polyadenylation signal (AATAAA) and the termination codon in the 5′-untranslated region which is in the same reading frame as the initiation codon. The nucleic acid sequences which flank the *Eco*RI site in CL-1 are shown aligned with the corresponding *Eco*RI site in pCB12. Those nucleotides which are different in pCB12 and CL-1 are indicated by an *.

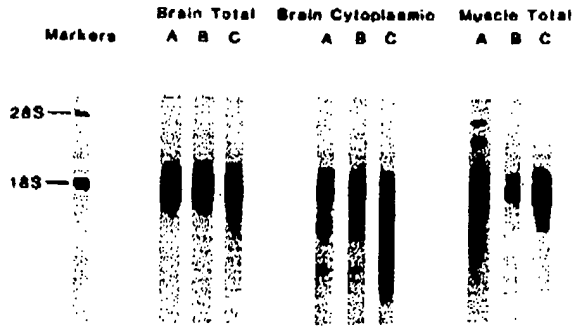

FIG. 3. **Analysis of the size and number of CaM mRNA sequences in chick brain and muscle.** Total poly(A⁺) RNA from chick brain and muscle and chick brain cytoplasmic poly(A⁺) RNA was prepared and resolved on 1.5% agarose gels as described under "Materials and Methods." The hybridization probes A, B, and C are described in Fig. 1. 1×10^7 cpm of ^{32}P-labeled probe was used for each hybridization filter. The size markers used were ^{32}P-labeled 18 S and 28 S rRNA from HeLa cells. The hybridization signals shown for total muscle RNA were obtained by an extended exposure time to compensate for the reduced degree of hybridization as compared to total brain RNA.

TABLE I

Interdomain nucleotide homology in chicken and eel CaM cDNAs

The assignment of domains is based on the amino acid sequence of bovine brain CaM (28). The domains are: I, amino acids 8–40; II, amino acids 44–76; III, amino acids 81–113; IV, amino acids 117–148. Overall sequence homology between amino acid-coding regions in chicken and eel cDNAs is 79%.

Domains	Per cent homology	
	Chicken/chicken	Chicken/eel
I/III	59	58
II/IV	64	56
I/II	45	41
II/III	51	44
III/IV	54	49
I/IV	50	43

TABLE II

Intradomain nucleotide homology between calcium-binding domains and subdomains in chicken and eel CaM cDNAs

The domains are detailed in the legend to Table I. The calcium-binding subdomains are as follows: I, amino acids 20–31; II, amino acids 56–67; III, amino acids 93–104; IV, amino acids 129–140.

Domain	Per cent homology	
	Domain	Subdomain
I	76	86
II	84	94
III	78	91
IV	75	91

*Hin*dIII-digested DNA (Fig. 4, *lanes H*) gave a slightly more complex pattern. Probe A′ hybridized to a fragment of 1.9 kb, probe B′ hybridized to fragments of 1.9 and 1.7 kb, while probe C′ hybridized to only the 1.7-kb fragment. From the hybridization patterns shown in Fig. 4, we can state that CaM is not represented by multiple genes in the chicken genome and predict that this gene is probably unique and has at least one intervening sequence which contains a *Hin*dIII site.

Isolation and Partial Characterization of a Genomic Clone for CaM—The cDNA insert from pCB12 and pCB15 were mixed and used as hybridization probes to screen a chicken genomic library constructed in the λ phage Charon 4A (kindly provided by T. Maniatis, Harvard University). A positive clone containing a 13.5-kb DNA insert was selected and designated CL-1. CL-1 has been characterized by partial DNA sequencing extending both directions from the unique *Eco*RI site located near the middle of the clone (see Fig. 5). The nucleotides which code for 30 amino acids just distal to

the *Eco*RI site (amino acids 11–40) are identical with those in the corresponding region of the chicken cDNA (see Fig. 2). This confirms the contention that CL-1 is derived from the chromosomal gene for CaM. Those nucleotides which are immediately 5′ to the *Eco*RI site do not code for any portion of the cDNA for CaM. Therefore, the *Eco*RI site must be at or within several nucleotides of the 3′-junction of the intervening sequence designated intron A in Fig. 5.

The unique *Eco*RI site in CL-1 divides the clone into 7.0-kb (5′-portion) and 6.5-kb (3′-portion) fragments. The 6.5-kb fragment was analyzed by endonuclease mapping and Southern hybridization and found to contain the structural sequences of the CaM gene which are 3′ from the *Eco*RI site in the cDNA. This structural sequence is interrupted by at least two introns. Intron B is characterized by the presence of an *Xba*I site 0.45 kb from the *Eco*RI site. This site could not be within the structural sequence because the cDNA sequence data reveal that the single *Xba*I site present is at least 0.7 kb from the *Eco*RI site. Therefore, the second *Xba*I site in the 6.5-kb fragment of CL-1 must reside in an exon. This assignment is supported by the distribution of *Dde*I, *Pst*I, and *Ava*II sites near this *Xba*I site which is identical with that found in the cDNA. The Southern hybridization analysis shown in Fig. 4 reveals an intron which is characterized by the presence of a *Hind*III site. Partial DNA sequencing of CL-1 from the *Eco*RI site (see Fig. 2) proves that the unique *Hind*III and *Eco*RI sites in the cDNA are not interrupted by an intron in CL-1. Therefore, the first *Hind*III in the 6.5-kb fragment of CL-1 resides in an exon while the second *Hind*III resides in an intron, intron C. Another exon region which is at least 3.8 kb 5′ from the *Eco*RI site has been identified by hybridization experiments using a pCB12 probe. The exact position of this exon is, as yet, unclear, and additional sequence analysis must be performed to determine if this exon contains the 5′ initiation site for the CaM gene.

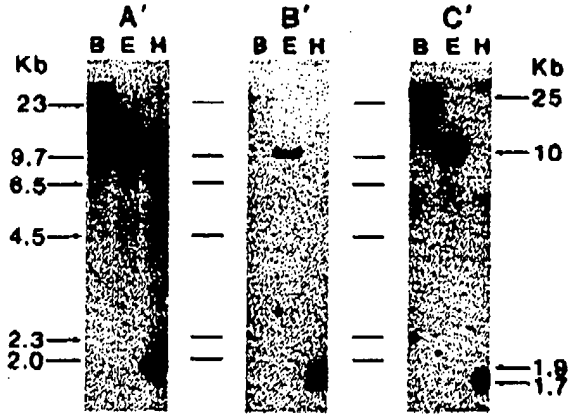

FIG. 4. Analysis of genome complexity for the CaM gene. 10 μg of chicken liver genomic DNA was digested with either *Bam*HI (*B*), *Eco*RI (*E*), or *Hind*III (*H*) and resolved on a 1% agarose gel. The DNA was transfered to nitrocellulose paper by the method of Southern (22) and hybridized to ^{32}P-labeled probes as described under "Materials and Methods." The probes *A′*, *B′*, and *C′* are described in Fig. 5. The *values listed on the left* indicate the position of size markers; the *values on the right* indicate the size of the various hybridization signals.

DISCUSSION

In this paper we describe the isolation and complete sequence of two cDNAs for chicken CaM which together represent a full length mRNA. This has allowed the determination of the amino acid sequence for chicken CaM and a nucleotide sequence comparison of cDNA for CaM from two different species. We also report the isolation and partial characterization of a genomic clone for chicken CaM which contains at least three intervening sequences.

As could be expected from the highly conserved nature of CaM from a variety of species, the primary amino acid sequence of chicken CaM (Fig. 2) is identical with that of human and bovine CaM (28, 29). Compared to eel CaM, there is a single conservative amino acid substitution at position 74 which is occupied by an Arg in the chicken and a Lys in the eel (7). The functional implications of this change are as yet unclear.

Consistent with the essentially identical primary amino acid sequence of chicken and eel CaM, the primary nucleotide sequences in the amino acid-coding portions of the respective cDNAs are highly homologous, with an overall homology of 79%. There are, however, regions in the chicken and eel cDNA molecules which have a degree of homology which is markedly higher than the overall homology. Specifically, the four pu-

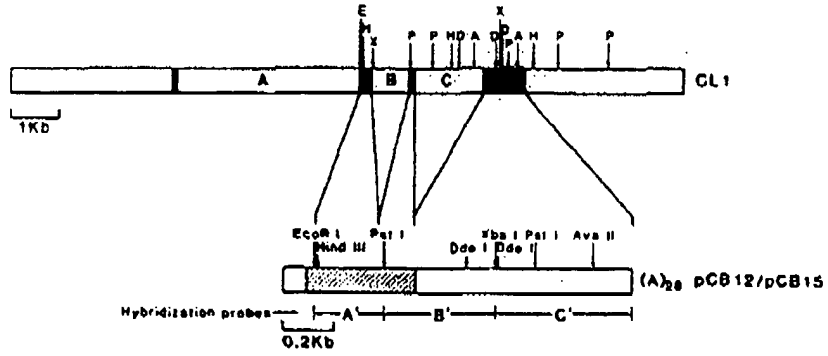

FIG. 5. Partial characterization of the genomic clone for CaM. The 6.5-kb fragment of CL-1 which extends 3′ from the *Eco*RI site to the end of the gene was subcloned and mapped with *Eco*RI, *E*; *Hind*III, *H*; *Pst*I, *P*; and *Xba*I, *X*. A 1.7-kb fragment, which contains the second *Xba*I site, was subcloned from a *Hind*III digest of the 6.5-kb fragment and mapped with *Ava*II and *Dde*I. The distribution of restriction sites in CL-1 is compared to those sites in pCB12/pCB15. The hybridization probes *A′*, *B′*, and *C′* were used in an analysis of genome complexity for the CaM gene. Probe A′ was obtained by digestion of pCB15 with *Eco*RI and *Pst*I, probes B′ and C′ obtained by digestion of pCB12 with *Pst*I and *Xba*I. The exons in CL-1 are indicated by the *black areas*; the introns are designated *A*, *B*, and *C*.

tative calcium-binding subdomains in the chicken and eel cDNAs have nucleotide homologies of 85–94% (Table II). The calcium-binding subdomains are each composed of 12 amino acids encoded by 36 nucleotides. We have compared adjacent segments of 36 nucleotides along the entire length of the amino acid-coding regions of the eel and chicken cDNAs and have found that the four calcium-binding subdomains have higher degrees of homology than all other portions of the molecule. This suggests that there is a greater pressure to conserve those sequences in the calcium-binding subdomains. In addition, the per cent homology between the four calcium-binding subdomains in the chick and eel cDNAs are roughly the same, particularly for subdomains II, III, and IV. This suggests that there is equivalent pressure to conserve these sequences.

In contrast to the amino acid coding portion, the 5′- and 3′-nontranslated regions of the cDNA for chicken and eel CaM have minimal homology (21 and 29%, respectively). Although very little direct sequence homology exists, the 3′-nontranslated regions in both the chicken and eel clones are AT-rich and contain numerous short stretches of dT and dA which is characteristic of other 3′-nontranslated regions (31–33). The 5′-nontranslated regions of the chicken and eel clones have distinctly different nucleotide compositions. The 5′-nontranslated region in the chicken contains 75% dC and dG while the eel contains only 50% dC and dG. The divergent nucleotide composition probably allows substantially different secondary structures in the respective 5′-nontranslated regions. This could be significant since the translation efficiency of mRNA is thought to be influenced by the secondary structure of the 5′-nontranslated region (34).

Fig. 3 demonstrates that hybridization probes prepared from pCB12 and pCB15 recognize two species of mRNA, with a major signal at 1600 nucleotides and a minor signal at 1900 nucleotides. Other tissues which have also been evaluated for their CaM mRNA complexity include gizzard, oviduct, liver, and cardiac muscle. In each case, the 1600-nucleotide and 1900-nucleotide mRNAs were observed but in concentrations which varied over a 10-fold range. There are at least three possibilities for the presence of two mRNAs. (*a*) One of the mRNA species represents contamination from a nuclear precursor. (*b*) The probes are recognizing an mRNA which does not code for CaM but codes for a protein which has a high degree of homology with CaM. (*c*) The mRNA species both code for CaM, the size difference being a result of differential processing of a primary RNA transcript.

The first possibility has been eliminated since hybridization of the probes to chicken brain cytoplasmic poly(A$^+$) RNA yield both hybridization signals. If one of the hybridization signals was due to contamination by a nuclear precursor then this signal should not be apparent after hybridization to cytoplasmic RNA. The second possibility is also unlikely. If the probes recognize an mRNA for a protein other than CaM, then the intensity of the signal should vary considerably depending on the nucleotide sequence of the hybridization probe. Specifically, probe C, which is derived entirely from the 3′-nontranslated region, should not give the same relative degree of hybridization to mRNAs for CaM and CaM-like proteins as does probe B which is derived almost exclusively from the amino acid-coding region of pCB12. What is observed is that all three probes hybridize to the 1900-nucleotide mRNA and, more importantly, the intensity of this signal relative to the intensity of the signal generated by the 1600-nucleotide mRNA is the same for all three probes. Therefore, we feel that both species of mRNA contain sequences which are complementary to all sequences present in pCB12 and pCB15. The 1600-nucleotide species is probably the parent mRNA for the cDNA described in this report while the 1900-nucleotide mRNA results from the third possibility listed above, that is, it is a product of differential processing of a primary RNA transcript. A similar mechanism has been found to exist for eel CaM (7) in which there are three species of mRNA which code for CaM, each of which results from differential recognition of multiple polyadenylation signals present in a common nuclear precursor. If this applies also to the chicken system, then we should be able to detect one additional polyadenylation signal with further characterization of the chicken genomic clone, CL-1.

The data presented in Figs. 4 and 5 reveal that only a single CaM gene with at least three introns exists in the chicken genome. This gene appears to be transcribed into the mRNA that is translated into the CaM found in all chicken tissues. Recently, we have screened the chicken genomic library with ^{32}P-labeled pCM109 under conditions of reduced stringency of hybridization and have isolated a clone which has 74% sequence homology to CL-1 and pCB12/pCB15. This gene, however, has no intervening sequences and encodes a protein which has 19 amino acid substitutions when compared to chicken CaM. The significance of this CaM-like gene remains to be determined.

CaM represents between 0.1 and 1.0% of the total protein present in every eukaryotic cell (3) and is involved in regulation of a wide variety of enzymes. In most instances CaM binds Ca^{2+}, undergoes a conformational change, and then interacts with the enzyme molecules (2, 3). However, CaM constitutes an integral subunit of phosphorylase kinase. One might have predicted that at least two CaM genes would exist, one to transcribe the mRNA used for "free" cytoplasmic CaM and another mRNA that would function to provide the CaM subunit of phosphorylase kinase. Since that prediction cannot be true, it is possible that one of the multiple species of CaM mRNA may contain information that would direct it to an intracellular area where the multisubunit enzyme is assembled. This possibility can be experimentally examined by comparing the abundance and number of CaM mRNA species present in ICR/IAn mice which are deficient in skeletal muscle phosphorylase and CaM (35) to those present in the wild type animal.

REFERENCES

1. Means, A. R. (1981) *Recent Prog. Horm. Res.* **37**, 333–367
2. Klee, C. B., and Vanaman, T. C. (1982) *Adv. Protein Chem.* **35**, 213–321
3. Means, A. R., Tash, J. S., and Chafouleas, J. G. (1982) *Physiol. Rev.* **62**, 1–38
4. Chafouleas, J. G., Pardue, R. L., Brinkley, B. R., Dedman, J. R., and Means, A. R. (1981) *Proc. Natl. Acad. Sci. U. S. A.* **78**, 996–1000
5. Chafouleas, J. G., Bolton, W. E., Hidaka, H., Boyd III, A. E., and Means, A. R. (1982) *Cell* **28**, 41–50
6. Munjaal, R. P., Chandra, T., Woo, S. L. C., Dedman, J. R., and Means, A. R. (1981) *Proc. Natl. Acad. Sci. U. S. A.* **78**, 2330–2334
7. Lagacé, L., Chandra, T., Woo, S. L. C., and Means, A. R. (1983) *J. Biol. Chem.* **258**, 1684–1688
8. Rosen, J. M., Woo, S. L. C., Holder, J. W., Means, A. R., and O'Malley, B. W. (1975) *Biochemistry* **14**, 69–78
9. Aviv, H., and Leder, P. (1972) *Proc. Natl. Acad. Sci. U. S. A.* **69**, 1408–1412
10. Bush, H. (1967) *Methods Enzymol.* **12**, 434–439
11. Tsai, S. Y., Roop, D. R., Tsai, M.-J., Stein, J. P., Means, A. R., and O'Malley, B. W. (1978) *Biochemistry* **17**, 5773–5780
12. Monahan, J. J., Harris, S. E., Woo, S. L. C., Robberson, D. L., and O'Malley, B. W. (1976) *Biochemistry* **15**, 223–233
13. McReynolds, L. A., Monahan, J. J., Bandura, D. W., Woo, S. L. C., Paddock, G. V., Salser, W., Dorson, J., Moses, R. W., and

O'Malley, B. W. (1977) *J. Biol. Chem.* **252**, 1840–1843
14. Chang, L. M. S., and Bollum, F. (1971) *J. Biol. Chem.* **246**, 909–916
15. Degert, M., and Ehrlich, S. D. (1979) *Gene* **6**, 23–27
16. Grunstein, M., and Hogness, D. S. (1975) *Proc. Natl. Acad. Sci. U. S. A.* **72**, 3961–3965
17. Denhardt, D. T. (1966) *Biochem. Biophys. Res. Commun.* **23**, 641–646
18. Manistis, T., Jeffrey, A., and Kleid, D. G. (1975) *Proc. Natl. Acad. Sci. U. S. A.* **72**, 1184–1188
19. Birnboim, H. C., and Doly, J. (1979) *Nucleic Acids Res.* **7**, 1513–1523
20. Katz, L., Williams, P. H., Sato, S., Leavitt, R. W., and Helinski, D. R. (1977) *Biochemistry* **16**, 1677–1683
21. Radloff, R., Bauer, W., and Vinograd, J. (1967) *Proc. Natl. Acad. Sci. U. S. A.* **57**, 1514–1521
22. Southern, E. (1975) *J. Mol. Biol.* **98**, 503–517
23. Thomas P. S. (1980) *Proc. Natl. Acad. Sci. U. S. A.* **77**, 5201–5205
24. Rave, N., Crkvenajakov, R., and Boedtker, H. (1979) *Nucleic Acids Res.* **6**, 3559–3567
25. Maxam, A. M., and Gilbert, W. (1977) *Proc. Natl. Acad. Sci. U. S. A.* **74**, 560–564
26. Maxam, A. M., and Gilbert, W. (1980) *Methods Enzymol.* **65**, 499–560
27. Sanger, F., and Coulson, A. R. (1978) *FEBS Lett.* **87**, 107–110
28. Watterson, D. M., Sharief, F., and Vanaman, T. C. (1980) *J. Biol. Chem.* **255**, 962–975
29. Sasagawa, T., Ericsson, L. H., Walsh, K. A., Schreiber, W. E., Fischer, E. H., and Titani, K. (1982) *Biochemistry* **21**, 2565–2569
30. Kretsinger, R. H., and Nockolds, C. E. (1973) *J. Biol. Chem.* **248**, 3313–3326
31. Proudfoot, N. J. (1976) *J. Mol. Biol.* **7**, 491–525
32. Benoist, C., O'Hare, K., Breathnach, R., and Chambon, P. (1980) *Nucleic Acids Res.* **8**, 127–141
33. McReynolds, L., O'Malley, B. W., Nisbet, A. D., Fothergill, J. E., Givol, D., Fields, S., Robertson, M., and Brownlee, G. G. (1978) *Nature (Lond.)* **273**, 723–729
34. Lee, K. A. W., Guertin, D., and Sonenberg, N. (1983) *J. Biol. Chem.* **258**, 707–710
35. Shenolikar, S., Cohen, P. T. W., Cohen, P., Nairn, A. C., and Perry, S. V. (1979) *Eur. J. Biochem.* **100**, 329–337